DA 5401
DATA ANALYTICS LAB

Let’s practice Machine Learning!

Course Objectives

01

Learn to transform a business problem into a **data science** problem

Learn to use Machine Learning to solve data-driven decision making.

02

Learn to evaluate the results of Machine Learning modeling

Learn to visualize and interpret the results

03

Learn to statistically analyze

Learn to perform model selection

Prerequisites

- DA5400 or any other equivalent courses on Machine Learning and/or Pattern Recognition
- Fluency in Python programming
- Ability to use third party libraries such as sklearn, numpy, pandas
- Linux or Mac or Windows WSL environment (Windows native discouraged)

The Process

- We will meet every week on Monday (P-slot)
- An hour of lecture will be given every week to introduce you to the practical aspects of an ML/Analytics topic.
- Students are expected to use the remaining time in the slot to do their assignments or clarify their queries.
- There will be at least eight assignments spread across the semester.
- At least three assignments will be evaluated in the class within the same evening.
- The remaining take-home assignments will be given a week for submission.
- The end-semester exam will be a Data Challenge competition.
- 75% weightage to the assignments, 20% to the data challenge (rank + report + viva) and 5% to class interaction.

References

- Sarah Guido, Andreas Muller., Introduction to Machine Learning through Python, 2024, PyStatMl — Statistics and Machine Learning in Python 0.5 documentation (duchesnay.github.io)
- Gareth James et al., Introduction to Statistical Learning, 2024, An Introduction to Statistical Learning (statlearning.com)

Git Repository

https://github.com/sudarsun/DA5401